

Supplementary material

Appendix 1: Significance and consistency of species responses to temperature and predictive performance of models, for all 25 species under study.

Appendix 2: Methods for the assessment of significance and consistency of temperature-only and all-predictors quantile regression models.

Appendix 3: Representation of temperature-only models and all-predictors models for all 25 species under study.

Appendix 4: Visual comparison of predictive performance of temperature-only and all-predictors models, respectively, for all 25 species under study.

Appendix 5: Maximum potential abundance predicted across the temperature sum gradient and its geographical projection in Finland for 1985 and 2040-2070 warming scenario, for 15 species with significant and consistent responses.

Appendix 6: Latitudinal shift of species maximum potential abundance predicted from 1985 to 2040-2070 in Finland, under IPCC scenario A1B.

Appendix 7: Partial effects of all environmental factors considered in the study on the maximum potential abundance of the 25 species.

Appendix 1: Significance and consistency of species responses to temperature and predictive performance of models, for all 25 plant species under study. *Temperature-only models (T-only)* are 95% quantile regression models including only effective temperature sum as predictor, while *all-predictors models (All-p)* include effective temperature sum, cumulative precipitation, soil texture, soil fertility, stand basal area and proportion of deciduous tree species as predictors. Prevalence: number of plots with positive abundance in 1985. L: measure of predictive performance based on the check loss function; lower values of L mean a better predictive capacity (see methods). Significance: whether the effect of effective temperature sum on maximum potential abundance is significant (x) or not (white space) in *T-only* and *All-p* models, considering 95% intervals (see Appendix S2 and S3). Consistency: whether the trend of the response of species maximum potential abundance to effective temperature sum is consistent (x) or not (white space) between *T-only* and *All-p* models, considering 95% intervals (see Appendix 2 and 3). Consistency of results is only analyzed when the response to effective temperature sum is significant in both *T-only* and *All-p* models. Species that show a significant response to effective temperature sum in both *T-only* and *All-p* models and a consistent response between them are selected for forecasting and highlighted in bold.

Specie	Plant group	Prevalence	L		Significance		Consistency
			T-only	All-p	T-only	All-p	
<i>Calamagrostis arundinacea</i>	Graminoid	227	0.304	0.279	X	X	X
<i>Calluna vulgaris</i>	Dwarf shrub	263	0.340	0.260			
<i>Carex globularis</i>	Graminoid	184	0.254	0.235	X	X	X
<i>Cladina arbuscula</i>	Lichen	294	0.314	0.217	X		
<i>Cladina rangiferina</i>	Lichen	354	0.290	0.213	X	X	X
<i>Deschampsia flexuosa</i>	Graminoid	642	0.260	0.248	X		
<i>Dicranum majus</i>	Bryophyte	159	0.287	0.235	X		
<i>Dicranum polysetum</i>	Bryophyte	610	0.216	0.218	X	X	X
<i>Dicranum scoparium</i>	Bryophyte	626	0.212	0.216	X	X	X
<i>Empetrum nigrum</i>	Dwarf shrub	261	0.304	0.254	X		
<i>Epilobium angustifolium</i>	Herb	224	0.288	0.252	X	X	X
<i>Hylocomium splendens</i>	Bryophyte	593	0.256	0.249	X	X	X
<i>Linnaea borealis</i>	Dwarf shrub	250	0.246	0.252			
<i>Luzula pilosa</i>	Graminoid	409	0.199	0.182		X	
<i>Maianthemum bifolium</i>	Herb	364	0.229	0.226	X	X	X
<i>Melampyrum pratense</i>	Herb	280	0.234	0.237	X		
<i>Pleurozium schreberi</i>	Bryophyte	796	0.177	0.168	X	X	X
<i>Pohlia nutans</i>	Bryophyte	274	0.221	0.218	X	X	X
<i>Polytrichum commune</i>	Bryophyte	414	0.307	0.312	X		
<i>Polytrichum juniperinum</i>	Bryophyte	249	0.285	0.270	X	X	X
<i>Solidago virgaurea</i>	Herb	210	0.234	0.230	X	X	X
<i>Trientalis europaea</i>	Herb	380	0.214	0.214	X	X	X
<i>Vaccinium myrtillus</i>	Dwarf shrub	754	0.198	0.173	X	X	
<i>Vaccinium uliginosum</i>	Dwarf shrub	202	0.247	0.247	X	X	X
<i>Vaccinium vitis-idaea</i>	Dwarf shrub	797	0.138	0.131	X	X	X

Appendix 2: Methods for the assessment of significance and consistency of *temperature-only* and *all-predictors* quantile regression models.

Response to temperature was considered insignificant, if it was possible to fit a horizontal line of constant response over the range of effective temperature sums included in our data within the estimated 95% confidence intervals. In such cases, our model does not provide statistically significant evidence against the null hypothesis of “no response to temperature”. Technically, this was checked by assessing, whether the maximum of the lower limits, $l(T)$, of 95% confidence intervals was smaller than the minimum of the upper limits $u(T)$ (Fig. A2-1a). Thus, response to temperature was considered significant if $l(T_+) > u(T_-)$, where T_+ is the temperature value that maximizes l and T_- the one that minimizes u . This implies that pairs of temperature sum values can be found, for which the confidence intervals of the predicted maximum potential abundance do not overlap (Fig. A2-1b).

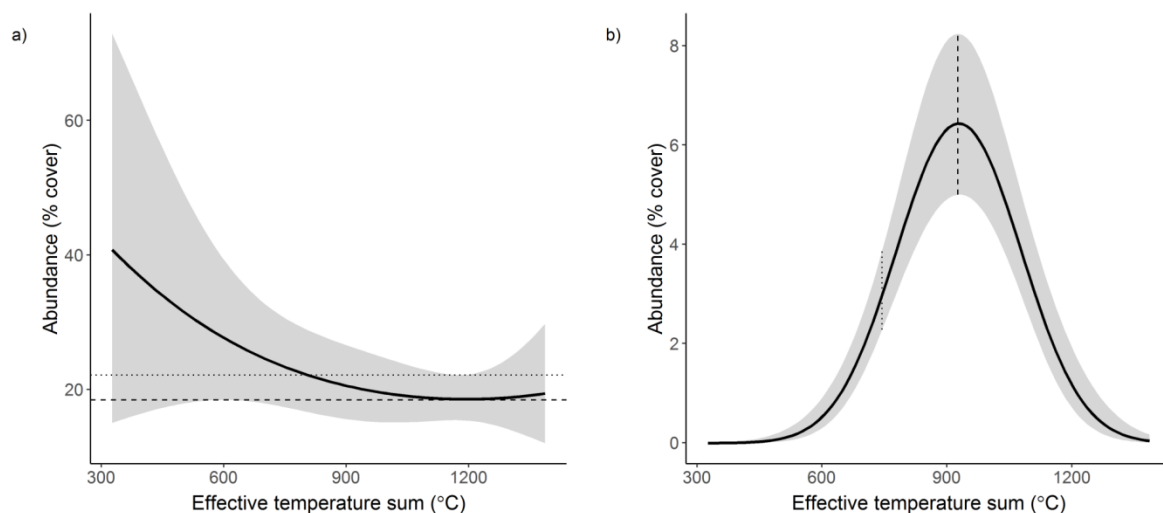


Figure A2-1. Examples of models with insignificant and significant response to temperature. a) *Temperature-only model* for *Calluna vulgaris*, illustrating maximum $l(T_+)$, of the lower limits of 95% confidence intervals (dashed line) and the minimum $u(T_-)$ of the upper limits (dotted line). Response of *C. vulgaris* to temperature was considered insignificant, because $u(T_-) > l(T_+)$. b) *Temperature-only model* for *Carex globularis* with significant response to temperature, illustrating a pair of non-overlapping confidence intervals (dashed and dotted line) associated with a pair of temperature sum values.

If the response to temperature was significant in both *temperature-only* and *all-predictors models*, we further checked whether the temperature response was consistent between the two models. This too was assessed on the basis of the estimated 95% confidence intervals by checking whether the pattern of high and low responses that was found significant in the *temperature-only model* can be reproduced within the confidence intervals of the all-predictors model. To specify how this was implemented, let us define T_+ as the value of effective temperature sum that maximizes $l_1(T)$, the lower limit of the confidence interval in the temperature-only model, and T_- as the value that minimizes $u_1(T)$. Furthermore, let L denote the set of temperature sum values T with “significantly low MPA”, $u_1(T) < l_1(T_+)$, and H the set with “significantly high MPA”, $l_1(T) < u_1(T_-)$. Then the responses were defined to be consistent, if the following conditions were satisfied by the confidence intervals $[l_2(T), u_2(T)]$ for the partial effect of effective temperature sum in all-predictors model (Fig. A2-2):

$$l_2(T) < u_2(T_+), \text{ for all } T \in L, \text{ and}$$

$$u_2(T) > l_2(T_-), \text{ for all } T \in H.$$

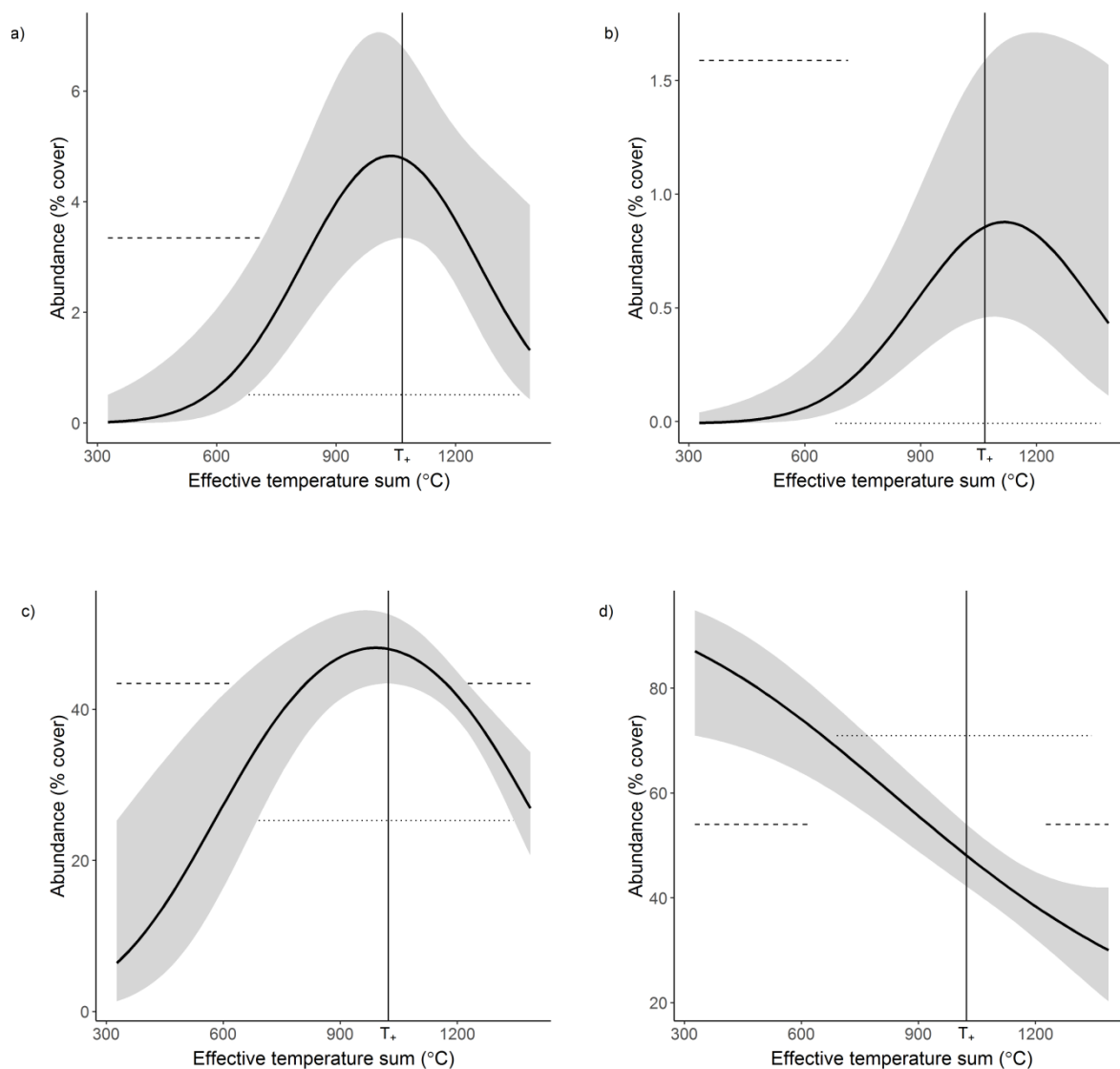


Figure A2-2. Examples with consistent and inconsistent response to temperature between *temperature-only* and *all-predictors models*. a) *Temperature-only model* for *Epilobium angustifolium*. T_+ maximizes $I_1(T)$, the lower limit of the 95% confidence interval. T_- , the value that minimizes $u_1(T)$ is the left-hand end of the range of effective temperature sums included in our data. The x-range of the dashed line segment shows the set L of temperature sum values with “significantly low MPA”; its y-level is the upper limit in the criterion $L=\{T: u_1(T) < I_1(T_+)\}$. Similarly, the dotted line segment illustrates H , the set with “significantly high MPA”. b) Partial effect of effective temperature sum in *all-predictors model* of *E. angustifolium*. The x-ranges of the dashed and dotted line segments are the same as in a), but their y-levels were determined from the *all-predictors model* to illustrate the limits $u_2(T_+)$ and $l_2(T_-)$ in our criteria for consistency. Models shown in a) and b) are consistent, because $l_2(T)$ is completely below the dashed line and $u_2(T)$ completely above the dotted line within the range of these segments. As a result, such a partial response fits within the confidence intervals of the *all-predictors model*, where MPA is lower at all $T \in L$ than at T_+ and greater at all $T \in H$ than at T_- . c) and d) are as a) and b), but for *Vaccinium*

myrtillus.; T_- is again the smallest effective temperature sum. The models shown in c) and d) are not consistent, because a response fitting within the confidence intervals can't have MPA lower than that at T_+ at any T in the lower part of L , and it can have MPA higher than that at T_- only in a very small part of H .

